What is claimed:

1. A method for speech recognition comprising:

a feature-amount extracting step for extracting a feature amount based on a frame of an input utterance;

a storing step for determining whether a current processing frame is within or at an end of a candidate word previously registered, and storing the candidate word on the basis of a first hypothesis-storage determining criterion when within a word and on the basis of a second hypothesis-storage determining criterion when at a word end;

a developing step for developing a hypothesis by extending utterance segments expressing the word when a stored candidate word is within a word and by joining a word to follow according to an inter-word connection rule when at a word end;

an operating step of computing a similarity of between the feature amount extracted from the input utterance and a frame-based feature amount of an acoustic model of the developed hypothesis, and calculating a new recognition score from the similarity and a recognition score of the hypothesis of up to an immediately preceding frame calculated from the similarity; and

a step of repeating the storing step, the developing step and the operating step until the processing frame becomes a last frame of the input utterance, and outputting, as a recognition result approximate to the input utterance, at least one of

18

hypotheses in the order of higher recognition score due to processing the last frame.

2. A method for speech recognition according to claim 1, wherein the first hypothesis-storage determining criterion is to select candidate words of within a predetermined threshold from a maximum value of the recognition score while the second hypothesis-storage determining criterion is to select a predetermined number of candidate words as counted from a candidate word maximum in the recognition score.

3. An apparatus for speech recognition comprising:

a feature-amount extracting section for extracting a feature amount based on a frame of an input utterance;

a search control section for controlling to develop a hypothesis by extending based on utterance segments to express a word when the hypothesis is within a word and by joining a word to follow according to an inter-word connection rule previously determined when at a word end;

a similarity computing section for computing a similarity of between a frame feature amount extracted from the input utterance and a frame feature amount of an acoustic model of the developed hypothesis;

a search operating section for operating a recognition score from the similarity and recognition score of the hypothesis of up to an immediately preceding frame;

19

a hypothesis determining section for determining whether a current processing frame is within a word or at a word end of the hypothesis and using the recognition score to select a candidate word according to a first determining criterion when within a word and to select a candidate word according to a second determining criterion when at a word end;

a hypothesis storing device for storing a hypothesis determined to be stored;

a word hypothesis registering device for registering as a new hypothesis the hypothesis and the recognition score; and

a recognition result output section for continuing the frame-based process to a last of the input utterance and outputting at least one hypothesis in the order of higher recognition score.

4. An apparatus for speech recognition according to claim 3, wherein the first determining criterion is to select candidate words of within a predetermined threshold from a maximum value of the recognition score while the second determining criterion is to select a predetermined number of candidate words as counted from a candidate word maximum in the recognition score.

5. A program for executing:

a feature-amount extracting step for extracting a feature amount based on a frame of an input utterance;

a storing step for determining whether a current processing frame is within or at an end of a candidate word previously registered, and storing the candidate word on the basis of a first hypothesis-storage determining criterion when within a word and on the basis of a second hypothesis-storage determining criterion when at a word end;

a developing step for developing a hypothesis by extending utterance segments expressing the word when a stored candidate word is within a word and by joining a word to follow according to an inter-word connection rule when at a word end;

an operating step of computing a similarity of between the feature amount extracted from the input utterance and a frame-based feature amount of an acoustic model of the developed hypothesis, and calculating a new recognition score from the similarity and a recognition score of the hypothesis of up to an immediately preceding frame calculated from the similarity; and

a step of repeating the storing step, the developing step and the operating step until the processing frame becomes a last frame of the input utterance, and outputting, as a recognition result approximate to the input utterance, at least one of hypotheses in the order of higher recognition score due to processing the last frame.

6. A program according to claim 5, wherein the first hypothesis-storage determining criterion is to select

candidate words of within a predetermined threshold from a maximum value of the recognition score while the second hypothesis-storage determining criterion is to select a predetermined number of candidate words as counted from a candidate word maximum in the recognition score.

7. A computer-readable recording medium recording a program for executing:

a feature-amount extracting step for extracting a feature amount based on a frame of an input utterance;

a storing step for determining whether a current processing frame is within or at an end of a candidate word previously registered, and storing the candidate word on the basis of a first hypothesis-storage determining criterion when within a word and on the basis of a second hypothesis-storage determining criterion when at a word end;

a developing step for developing a hypothesis by extending utterance segments expressing the word when a stored candidate word is within a word and by joining a word to follow according to an inter-word connection rule when at a word end;

an operating step of computing a similarity of between the feature amount extracted from the input utterance and a frame-based feature amount of an acoustic model of the developed hypothesis, and calculating a new recognition score from the similarity and a recognition score of the hypothesis of up to

an immediately preceding frame calculated from the similarity; and

a step of repeating the storing step, the developing step and the operating step until the processing frame becomes a last frame of the input utterance, and outputting, as a recognition result approximate to the input utterance, at least one of hypotheses in the order of higher recognition score due to processing the last frame.

8. A computer-readable recording medium recording a program according to claim 7, wherein the first hypothesis-storage determining criterion is to select candidate words of within a predetermined threshold from a maximum value of the recognition score while the second hypothesis-storage determining criterion is to select a predetermined number of candidate words as counted from a candidate word maximum in the recognition score.

23